6-12-03

REMARKS

The Examiner rejected Claims 1-12 under 35 U.S.C. 103(a) as being unpatentable over Yu, US Patent No. 6,351,775 and further in view of Desai, US Patent 6,434,608. Applicant traverses this rejection. To sustain a rejection under 35 U.S.C. 103, the Examiner must show that the combined references teach each of the elements of the claim or that there is some motivation in the art for altering one of the teachings to arrive at the combined set of teachings. "The mere fact that a reference could be modified to produce the patented invention would not make the modification obvious unless it is suggested by the prior art." (*Libbey-Owens-Ford v. BOC Group*, 4 USPQ 2d 1097, 1103). In addition, the Examiner must show that there is some motivation in the art that would cause someone of ordinary skill to combine the references, and that in making the combination, there was a reasonable expectation of success. Where the claimed subject matter has been rejected as obvious in view of a combination of prior art references, a proper analysis under section 103 requires, *inter alia*, consideration of two factors: (1) whether the prior art would have suggested to those of ordinary skill in the art that they should make the claimed composition or device, or carry out the claimed process; and (2) whether the prior art would also have revealed that in so making or carrying out, those of ordinary skill would have a reasonable expectation of success... Both the suggestion and the reasonable expectation of success must be founded in the prior art, not in the applicant's disclosure. *In re Vaeck*, 20 USPQ2d 1438, 1442(CAFC 1991).

In making this rejection, the Examiner stated that in regards to Claim 1, Yu discloses a method for operating a server cluster comprising N server nodes to service client requests, each client request being directed to one of a plurality of sites hosted on said server cluster, each site being identified by a domain name and each server node being identified by an address on a network connecting said clients to said server nodes. The Examiner points to the abstract of Yu in support of this assertion, (Abstract). Applicant must disagree with the Examiner's reading of Yu.

Yu teaches grouping objects specified by URLs into classes. Yu teaches a system in which URLs are grouped into classes and the various classes are assigned to the various

servers. Yu defines the term URL as the unique address of information on the Internet (Column 1, line 56-Column 2, line 3). An example of a URL according to Yu is "http://www.philipyu.com:80/table.html". The URL in this example includes the domain name "philipyu.com" as a part thereof; however, a URL, in general, need not include a domain name since not all IP addresses have domain names, e.g. http://186.197.3.35/table.html.

In contrast, the present invention is directed to the grouping of sites on servers. A site is a collection of files associated with a single domain name. Hence, a site according to the present invention may be viewed as a class according to Yu in which all of the URLs refer to the same domain name. Yu does not teach forming such a class. Yu teaches only forming collections of URLs, referred to as "classes", and making assignments of those classes to servers (column 6, lines 18-29). There is no teaching in Yu that the selected URLs that form a class all refer to the same domain name. In fact, the examples given starting at Column 6, line 18 utilize a hash function to divide the URLs into classes irrespective of any domain name contained in the URLs. Yu also refers to the use of the PICS labels for dividing URLs into classes by content. Since the PICS labels do not depend on the domain name, this mapping of URLs into classes does not depend on the domain name associated with the URLs. Hence, in general, a class according to Yu will include URLs belonging to a plurality of sites and some that do not belong to any site having a domain name. That is, a class as used by Yu is not identified by "a domain name". In contrast, a site according to the present invention is identified by a single domain name.

Claim 1 requires the measurement of the computational resources required to service the requests to each of the sites over a first period of time. The Examiner admits that Yu does not teach this limitation. The Examiner looks to Desai for this missing teaching. According to the Examiner, Desai teaches the tracking of requested objects that are retrieved from servers and sent to the requesting host and monitoring how many cache misses have occurred in supplying such objects. The Examiner specifically points to lines 14-44 of column 3. According to the Examiner, an object is equivalent to a computational resource. First, the object being returned is the file specified by the URL; hence, the object is not a computational resource. Second, the cited passage refers to recording the identity of objects for which a request has generated a cache miss in a table. It should be noted that if all of the objects on

the server were stored in the cache, the system taught in Desai would not store any information. Hence, the table identified by the Examiner is not a measure of the computational resources needed to service the requests of each of the sites. In fact, it is not even a measure of the computational resources needed to service a request for a particular object.

In this regard, it should also be noted that there is no correlation between a cache miss and the computational resources needed to deliver the object specified by the URL. A cache miss is generated when the object in question is not in the cache. This occurs the first time the object is requested, regardless of the computational resources required to supply the object. In addition, a miss occurs on subsequent requests for the object if the object has been displaced from the cache by another object. In this case, the miss is a function of the sequence of requests for objects, not the computational resources needed to service the requests for the objects.

The Examiner also points to the passage at column 6, line 62 to column 7, line 15 as supporting the Examiner's contention that Desai teaches measuring computational resources. Applicant disagrees with the Examiner's reading of this passage. The cited passage refers to an algorithm for adjusting a threshold value that determines how may cache misses must occur for an object before the object is cached. The passage refers to keeping track of cache hits per object requested. As noted above, there is no correlation between cache misses, or hits, and the computational resources needed to deliver an object in response to a request for that object.

The Examiner states that one would combine the teachings of Desai into the system of Yu to determine how many cache misses to require before caching an object. Note that this combined system does not satisfy the limitations of Claim 1. The combined system would count the number of cache misses generated by an object that is not currently cached to determine when the object is to be cached on the server to which it is assigned. That is, the resultant system is that of Yu with a more efficient caching algorithm on the individual servers. This is not a system in which the grouping of URLs on the servers is adjusted to keep the difference in the sum of cache misses generated by each site within a predetermined error value. In fact, this system operates independent of the grouping of URLs on the servers.

According to the Examiner, Yu teaches grouping said sites into N groups, each group being assigned to a corresponding one of said server nodes such that for each pair of groups, the difference in the sum of said measured computational resources is within a first predetermined error value (col. 6, lines 31-36; col. 7, lines 18-37; and col. 9, lines 27-33). The Examiner refers to the teaching of SA(j) as the number of requests for object classes assigned to server j, but is silent as to how this relates to this contention. Since the Examiner has already admitted that Yu does not teach measuring computational resources, Applicant cannot see how the Examiner has arrived at this conclusion. At most, Yu teaches grouping classes, not sites, utilizing the number of requests C(J,I) for objects in class I on server J. The quantity SA(J) is the sum of C(J,I) over the classes. The Examiner refers to column 6, lines 31-36, column 7, lines 18-37, and column 9, lines 27-33 in support of this assertion. The first passage merely refers to monitoring the total load on each server and changing the class to server assignments to balance the load. The second passage refers to how a change in assignment is communicated once such a change has been determined to be advantageous. The third passage refers to keeping statistics on the number of requests for objects in class I on server J.

The algorithm taught in Yu moves a class as a whole from one server to another if the servers are unbalanced as measured by the number of requests for each class on each server. It should be noted that Yu measures only the number of requests for each class on each server, not the number of requests for each site on each server. Hence, even if one were to substitute measuring the number of cache misses for each class on each server for C(J,I), and the quantities derived therefrom, i.e., SA(J), the resultant modified version of Yu would not satisfy this limitation of Claim 1. In this regard, it should also be noted that the same argument applies to the limitation of "measuring the computational resources required to service said requests to each of said sites over a first time period", since, at most, the modified version of Yu would not measure any quantity related to each of the sites, no less the computational resources needed. Accordingly, Applicant submits that the Examiner has not made a *primia facia* case for obviousness with respect to Claim 1 or the claims dependent therefrom.

With reference to Claim 3, the Examiner maintains that Desai teaches recording information identifying each returned file, the size of that file, and the number of times that file was returned. The Examiner points to Desai, column 3, lines 14-44 and column 6 line 62 to column 7, line 15 in support of this assertion. First, as noted above, Desai only teaches recording the identity of the files for which a cache miss has occurred and the number of cache misses for those files prior to the files being cached. Hence, the recorded information does not identify **each returned file.** Furthermore, the number of cache misses is not a measure of the number of times the file was returned, since once the file is cached, Desai no longer keeps count of the number of times the file is requested. Finally, there is no teaching of recording the file size for these files, no less all returned files. Hence, Applicant submits that there are additional grounds for allowing Claim 3 and the claims dependent therefrom.

With reference to Claim 4, the Examiner maintains that Yu-Desai disclose the limitation that each of said server nodes comprises a cache memory for facilitating the return of said files in response to said request and wherein said step of grouping said sites also depends on the amount of memory in said cache memory on each of said servers. The Examiner refers to Yu, column 8, line 10 to column 9, line 10 as supporting this contention. The cited passage refers to the manner in which the cache manager on each of the servers decides whether or not an object is to be cached on that server. Similarly, Desai also uses the size of the cache to determine whether or not an object should be cached. The Examiner also points to step 1040 shown in Figure 10 of Yu. Applicant must respectfully point out that the step refers to computing the average number of requests for objects per server. The term "M" refers to the number of servers, not memory. Hence, neither reference teaches grouping objects onto servers based on the amount of memory in the cache on each server, no less grouping sites in this manner. Accordingly, there are additional grounds for allowing Claim 4 and the claims dependent therefrom.

As to Claim 5, the Examiner maintains that Yu-Desai disclose the limitation that said groups are chosen such that said files returned during said first time period more than a predetermined number of times can be stored simultaneously in said cache memory. The Examiner points to Desai, column 3, lines 14-44 in support of this contention. Applicant must disagree. The cited passag refers to the manner in which the system determines whether or not an object is to be cached on a server. There is no teaching of grouping objects,

no less sites, on servers such that all files returned more than a predetermined number of times can be stored simultaneously in the cache memory. Hence, there are additional grounds for allowing Claim 5.

As to Claim 6, the Examiner stated that Yu-Desai discloses the limitation that said measurement of said computational resources further comprises measuring the number of bytes of data returned in response to said requests for each site during said first time period. The Examiner refers to Desai, column 6, line 62 to column 7, line 10 as supporting this contention. The cited passage refers to using the size of the object relative to an object threshold to determine if the object should be cached in a network cache. There is no teaching of measuring the number of bytes of data returned in response to each request during a first time period. Furthermore, there is no teaching of grouping objects, no less sites, based on such a measurement. The reference uses the information to determine if an object is to be cached, not to determine on which server the object is to be located. Further, once the object is cached, all data collection for that object is suspended. Hence, there are additional grounds for allowing Claim 6 and the claims dependent therefrom.

As per Claim 7, the Examiner stated that Yu-Desai teach the limitation of estimating the number of bytes of data returned directly from said cache memory in servicing said requests for each site during said first time period. The Examiner points to Desai, column 6, line 62 – column 7, line 10 as supporting this contention. As noted above, the cited passage teaches using the size of a file relative to an object size threshold to decide whether to cache the file. There is no teaching of estimating the number of bytes returned directly from the cache memory in servicing requests. In this regard, it should be noted that once an object is cached, there is no teaching that the system uses the size of the object for any purpose, no less making such an estimate. Accordingly there are additional grounds for allowing Claim 7.

As to Claim 11, the Examiner stated that Yu-Desai disclose measuring the computational resources required to service said requests to each of the sites over a second time period (Desai, col. 3, lines 14-44), and grouping said sites into N new groups, by swapping sites between said previous groups, each new group being assigned to a corresponding one of said server nodes such that for each pair of new groups, the difference in the sum of said measured computational resources over said second time period is within a

second predetermined error value (Yu, col. 9, line 51 – col. 10, line 42). Applicant respectfully disagrees. As noted above, Yu teaches swapping classes, not sites. Since a site may be split between a number of classes, moving a class does not necessarily imply moving even one site. Accordingly, there are additional grounds for allowing Claim 11 and the claim dependent therefrom.

As to Claim 12, the Examiner stated that Yu-Desai disclose wherein said new groups differ from said previous groups by as few site swaps as possible (Yu, col. 10, lines 25-42). Applicant disagrees. As noted above, Yu teaches moving classes not sites. Since a class can have a plurality of sites, moving a class does not minimize the number of sites that are moved. For example, if one site has an object that is being requested a large number of times, the minimum move would be to move just the site containing that object. In the scheme taught in Yu, one would move the class containing the object. As noted above, this may involve moving a site, a group of sites, or no site at all. Accordingly, there are additional grounds for allowing Claim 12.

I hereby certify that this paper is being sent by FAX to 703-746-7238.

Respectfully Submitted,

Calvin B. Ward
Registration No. 30,896

Date:June 12, 2003
Hewlett-Packard Co.
Intellectual Property Administration
P.O. Box 272400
Fort Collins, CO 80527-2400
Telephone (925) 855-0413
Telefax (925) 855-9214